



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

출원 번호 : 10-2003-0009607
Application Number

출원 년 월 일 : 2003년 02월 15일
Date of Application FEB 15, 2003

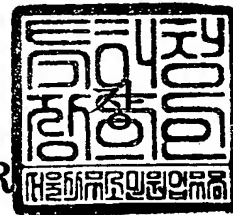
출원인 : 삼성전자주식회사
Applicant(s) SAMSUNG ELECTRONICS CO., LTD.



2003 년 08 월 08 일

특 허 청

COMMISSIONER





1020030009607

출력 일자: 2003/8/9

【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0008
【제출일자】	2003.02.15
【국제특허분류】	G10L
【발명의 명칭】	오디오 데이터 인코딩 장치 및 방법
【발명의 영문명칭】	Audio data encoding apparatus and method
【출원인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	2003-003435-0
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2003-003436-7
【발명자】	
【성명의 국문표기】	장흥엽
【성명의 영문표기】	JANG,Heung Yeop
【주민등록번호】	710810-1019629
【우편번호】	442-811
【주소】	경기도 수원시 팔달구 영통동 963-2 신나무실 쌍용아파트 542동 101 호
【국적】	KR
【발명자】	
【성명의 국문표기】	김병일
【성명의 영문표기】	KIM,Byeong Il
【주민등록번호】	750721-1357117



1020030009607

출력 일자: 2003/8/9

【우편번호】	156-756
【주소】	서울특별시 동작구 흑석1동 중앙대학교
【국적】	KR
【발명자】	
【성명의 국문표기】	장태규
【성명의 영문표기】	JANG,Tae Kyu
【주민등록번호】	551113-1052810
【우편번호】	156-756
【주소】	서울특별시 동작구 흑석1동 중앙대학교
【국적】	KR
【심사청구】	청구
【취지】	특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인 이영필 (인) 대리인 이해영 (인)
【수수료】	
【기본출원료】	20 면 29,000 원
【가산출원료】	7 면 7,000 원
【우선권주장료】	0 건 0 원
【심사청구료】	13 항 525,000 원
【합계】	561,000 원
【첨부서류】	1. 요약서·명세서(도면)_1통

【요약서】**【요약】**

본 발명은 적은 계산량으로 오디오 데이터를 인코딩하는 장치 및 방법에 관한 것이다. 본 발명의 오디오 데이터 인코딩 장치는 시간영역의 오디오 신호를 입력받아 주파수 영역의 신호로 변환하는 시간/주파수 매핑부; 상기 변환된 주파수 영역의 오디오 신호를 입력받아 인코딩하고자 하는 오디오 포맷에 대응되는 스펙트럴 처리를 수행하는 스펙트럴 처리부; 상기 변환된 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 마스킹 임계치 계산부; 및 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 공통이득을 조절하여 양자화 잡음 곡선을 소정의 에너지 분포 곡선에 정합시키는 양자화 잡음 곡선 조절부를 포함한다. 본 발명의 인코딩 장치는 심리음향 모델을 직접 사용하지 않고 주파수별 에너지 분포를 통하여 임계잡음 레벨의 상대적인 대역별 분포 형태와 비슷한 분포를 산출함으로써 쉽게 구현될 수 있다.

【대표도】

도 3

【명세서】

【발명의 명칭】

오디오 데이터 인코딩 장치 및 방법{Audio data encoding apparatus and method}

【도면의 간단한 설명】

도 1은 종래의 오디오 인코더의 블록도이다.

도 2a 내지 도 2b는 마스킹 효과를 설명하기 위한 도면이다.

도 3은 본 발명의 오디오 인코딩 장치의 블록도이다.

도 4a 내지 도 4d는 스케일 팩터 밴드의 에너지를 근사화하는 과정을 설명하기 위한 도면이다.

도 5는 본 발명의 오디오 인코딩 방법의 플로우차트이다.

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

<6> 본 발명은 오디오 데이터의 인코딩에 관한 것으로, 구체적으로는 적은 계산량으로 오디오 데이터를 인코딩하는 장치 및 방법에 관한 것이다.

<7> 오디오 데이터를 소정의 형식으로 압축하는 인코더는 심리음향모델을 사용하고, 심리음향모델에서 수행되는 계산결과를 바탕으로 주파수 밴드별 양자화 잡음을 다단계 제어 루프에 의해서 조절한다. 여기서 양자화(quantization)는, 샘플링된 신호값을 일정한 대표값으로 나타내기 위하여 스텝모양의 정수값으로 표현하는 것으로 이 과정에서 양자화 잡음이 발생한다. 원래신호와 양자화된 신호와의 오차성분인 양자화 잡음은 양자화에

사용되는 비트수가 많아질수록 작아진다. 동영상 및 음성에 대한 압축 표준인 엠펙(MPEG)에서는, DCT(Discrete Cosine Transform) 또는 MDCT(Modified Discrete Cosine Transform) 변환에 의해 계산된 계수를 어떤 값으로 나누어 작은 값의 계수로 표현함으로써 부호화량을 감소시킨다.

<8> 그리고, 상술한 다단계 제어루프란, 종래의 양자화 잡음 분포 조절 방법에서 사용되는 것으로, 모든 주파수 대역에 공통으로 적용되는 공통 이득(common gain)을 조정하여 정해진 비트율에 비트 사용량을 맞추는 내부 루프와, 각 주파수 대역별로 양자화 잡음의 크기를 조정할 수 있는 스케일팩터 밴드 이득을 조정하는 외부 루프를 의미한다. 내부 루프에서는 각 주파수 대역별로 조정된 스케일팩터 밴드 이득을 적용해 부호화하여 사용된 비트량을 합산하여 이 값이 소정의 허용된 값을 초과하는 경우에는 공통 이득을 증가시켜 비트 사용량을 허용치 이하로 만들고, 외부 루프에서는 각 주파수 대역별로 주어진 임계치를 넘지 않도록 주파수 대역별 스케일팩터 밴드 이득을 일정한 크기로 증가시킨다. 이러한 과정을 반복적으로 수행하여 모든 주파수 대역에서의 양자화 잡음이 임계치를 넘지 않을 때까지 계속한다.

<9> 일반적으로 오디오 데이터를 인코딩하는 것이 디코딩하는 것보다 10배 이상의 계산량을 요구하는데, 이중 심리음향모델에서의 FFT 수행, 톤리티(tonality) 계산, 마스크 임계치(mask threshold)의 계산 및 프레임간의 프로세싱 등의 수행이 전체 계산량의 50% 정도를 차지하고, 비트율과 노이즈를 제어하는 다단계 제어 루프의 수행이 전체 계산량의 40% 정도를 차지하여, 인코더가 복잡해지는 원인이 된다.

<10> 도 1은 종래의 오디오 인코더의 블록도이다.

- <11> 오디오 인코더는 시간/주파수 매핑부(110), 스펙트럴 처리부(120), 양자화부(130), 심리음향모델(140), 비트 할당부(150) 및 비트스트림 생성부(160)를 포함한다.
- <12> 시간/주파수 매핑부(110)는 시간영역의 PCM(Pulse Code Modulation) 오디오 데이터를 입력받아 주파수 영역의 신호로 변환한다. 인코딩하는 포맷에 따라서 시간/주파수 매핑부(110)에서 수행되는 처리가 달라지는데, AAC(Advanced Audio Coding) 포맷 또는 MP3(MPEG-1 layer 3) 포맷으로 인코딩할 때는 MDCT(Modified Discrete Cosine Transform)가 수행된다.
- <13> 스펙트럴 처리부(120)는 주파수 영역의 신호를, 인코딩하는 오디오 포맷에 맞는 스펙트럴 처리를 수행한다. 이러한 스펙트럴 처리의 예로는, TNS(Temporal Noise Shaping), LTP(Long Term Prediction), PNS(Perceptual Noise Substitution), I/C, M/S 등이 있다. 양자화부(130)는 스펙트럴 처리된 주파수 영역의 오디오 데이터에 대하여 양자화를 수행한다.
- <14> 심리음향모델(140)은 FFT 수행부(141)와 마스킹 임계치 계산부(142)를 포함하며, 주파수 영역에서의 인간의 청각 특성을 반영한다. 심리음향모델(140)에서 수행되는 처리는 후술한다. 이제, 주파수 영역에서의 인간의 청각 특성을 도 2a 내지 도 2b를 참조하여 설명한다.
- <15> 도 2a 내지 도 2b는 마스킹 효과를 설명하기 위한 도면이다.
- <16> 도 2a에 도시한 바와 같이 소정의 음압을 가진 오디오 신호 A(210)가 존재할 때, 음압이 오디오 신호 A(210)의 음압보다 어느 정도 작은 소리(220)는 들리지 않게 되는데, 이렇게 특정 오디오 신호에 대해서 가청 주파수 내에서 인간이 들을 수 있는 최소한의

음압 레벨의 곡선을 마스킹 곡선(230)이라고 한다. 따라서, 오디오 신호 B(220)는 마스킹 곡선(230)보다 음압이 작으므로 인간의 귀로 들을 수 없고, 오디오 신호 C(240)는 마스킹 곡선(230)보다 음압이 크므로 인간의 귀로 들을 수 있다.

<17> 만일 여러개의 피크치(250, 260, 270)가 도 2b에 도시한 바와 같이 위치한다면, 각각의 피크치에 대한 마스킹 곡선(251, 261, 271)이 존재하고, 이들 마스킹 곡선을 연결하면 전체적인 마스킹 곡선을 얻을 수 있다.

<18> 이와 같이 인간의 귀로 들을 수 있는 주파수를 일정간격으로 나누어, 마스킹 임계치 이상의 음압을 가진 오디오 데이터만을 양자화하는 것을 심리음향모델(psychoacoustic model)을 사용한 양자화라고 하고, 엠팩(MPEG)과 같은 압축방법에서 사용된다. 그러나, 64Kbps 이하의 저속의 비트율로 오디오 신호를 압축하는 경우에는 양자화시에 사용될 수 있는 비트의 수에 한계가 있기 때문에 MPEG 표준에서 제시하고 있는 일반적인 오디오 압축방법은 효과적으로 오디오 신호를 압축하는데 적합하지 않다.

<19> 비트 할당부(150)는 심리음향모델(140)에서 계산된 결과를 입력받아 비트 할당을 수행한다. 그리고, 양자화된 오디오 데이터를 소정의 형식에 맞게 팩킹하는 과정은 비트 스트림 생성부(160)에서 수행된다.

<20> 종래의 MPEG 오디오 인코딩 과정은 다음과 같다. MPEG 인코딩 알고리즘은 ISO/IEC 14496-3 표준에 상세히 설명되어 있다.

<21> 우선, 시간 영역의 신호를 주파수 영역의 신호로 변환하기 위해 PCM 오디오 데이터를 시간/주파수 매핑부(110)로 입력받는다. 그리고, 시간 영역의 PCM 오디오 데이터는 심리음향모델(140)로도 입력된다.

- <22> 그리고, 심리음향모델(140)은 인간의 주파수 영역에 대한 청각 특성을 반영하기 위하여 입력된 오디오 데이터를 FFT를 이용하여 주파수 영역의 데이터로 바꾸고, 인간의 공통된 청각 특성이 비슷한 임계대역(critical band)으로 나눈다. 특정 임계 대역에 신호가 존재하면 이웃하는 임계대역에 존재하는 신호성분을 인지할 수 있는 음압의 레벨이 올라가게 되는데(도 2a 내지 도 2b 참조), 이러한 청각 특성을 마스킹 특성(masking effect)라고 한다.
- <23> 다음으로, FFT에 의해 변환된 주파수 영역 오디오 데이터의 마스킹 특성을 사용하여 각 임계대역 별로 마스킹 임계치(masking threshold)를 산출한다. 이때 마스킹 특성을 고려하여, 해당 주파수의 오디오 데이터가 톤 성분인지 잡음 성분인지를 구별해야 한다. 잡음 성분이 톤 성분으로 선택되는 것을 방지하기 위해서 과거 두 블록의 주파수 성분을 가지고 선형 예측을 하여 톤 성분인가를 판단한다.
- <24> 시간 영역에서 한 블록의 신호 구간 내에 음압이 큰 신호와 음압이 아주 작은 신호가 같이 포함되어 있을 때, 주파수 변환 과정과 양자화 과정을 거치면 음압이 큰 신호의 양자화 잡음이 음압이 아주 작은 신호에 포함되어 잡음이 들리게 되는데, 이를 프리에코 현상(pre-echo effect)이라고 한다. 이러한 프리에코 현상을 방지하기 위해 한 블록에 대해, 긴 윈도우 블록(long window block)을 사용한 주파수 변환을 수행하는 대신에, 한 블록을 8구간으로 나눈 짧은 윈도우 블록(short window block)을 사용한 주파수 변환을 수행한다. 심리음향모델에서는 긴 윈도우 블록(long window block)과 짧은 윈도우 블록(short window block)을 선택하기 위해서 심리음향 엔트로피(perceptual entropy)를 계산한다.

- <25> 그리고 나서, 스펙트럴 처리부(120)는 오디오 데이터를 압축하기 위해 주파수 영역으로 표현된 신호 성분간의 잉여성분을 제거한다.
- <26> 주파수 영역의 신호 성분들은 스케일팩터 밴드(scalefactor band)로 구분되고 각 신호성분은 해당 스케일팩터 밴드내에서 공통으로 적용되는 이득과 양자화 값의 곱으로 나타내어진다. 이때 이득을 결정짓는 요소는 전체 주파수 밴드에 공통적인 값인 공통 게인(common gain)과 스케일팩터 밴드별로 구분되는 스케일팩터(scalefactor)가 있다. 공통 게인은 목표 비트율을 맞추기 위해 조절되는 값이고, 스케일팩터는 스케일팩터 밴드별로 양자화 잡음을 조절하기 위한 값이다. 스케일팩터 밴드별 허용되는 양자화 잡음은, 심리음향모델에서 산출한 마스킹 임계치를 이용하여 결정된다.
- <27> 이와 같이 종래의 오디오 인코딩 방법에서는, 심리음향모델에서 마스킹 임계치를 계산하기 위해서, 주파수 영역으로 변환을 위한 FFT 연산, 마스킹 특성을 적용하는 스프레딩(spreading) 함수의 처리, 프레임간 선형 예측을 통한 톤리티(tonality) 처리 등이 수행되어 많은 계산량을 요구한다. 그리고, 심리음향 모델에서의 FFT 연산과는 별도로 주파수 영역에서의 신호처리를 위해 시간 영역 신호에 대하여 DCT를 수행한다. 따라서, 인코더의 데이터 처리 시간을 크게 늘린다는 문제점이 있다. 즉, 기존 MPEG 오디오 압축에서는 고품질의 얻기 위한 노력으로 심리음향 모델을 사용하지만, 이에 따른 데이터의 복잡한 처리 과정과 연산량의 증가를 피할 수 없다는 문제점이 있다.
- <28> 그리고, 양자화 과정에서는 주파수 대역별 비트할당을 이용하여 양자화 잡음을 조절하는 과정과 전체 비트율을 맞추기 위한 과정이, 원하는 비트율을 맞추면서 허용 잡음치내에 들어올 때까지 반복하여 수행된다. 그러나 저비트율의 오디오 인코딩 과정에서는 블록별 사용할 수 있는 비트의 수가 적어 대역별 양자화 잡음을 심리음향모델에서 산출

한 허용할 수 있는 잡음의 크기보다 작게 되도록 만족시키지 못하고 양자화 과정을 종료한다는 문제점이 있다.

【발명이 이루고자 하는 기술적 과제】

<29> 본 발명이 이루고자 하는 기술적 과제는, 종래의 오디오 인코딩 수행시에 사용되는 복잡한 연산과정을 필요로 하는 심리음향모델을 사용하지 않고, 오디오 신호의 대역별 에너지 분포를 계산하여 상대적으로 적은 계산량을 사용하여 심리음향 모델을 추정하는 오디오 인코딩 장치 및 방법을 제공하는데 있다.

<30> 본 발명이 이루고자 하는 다른 기술적 과제는, 종래의 양자화 잡음 조절방법에서 사용되는, 비트율과 양자화 잡음 분포를 동시에 만족시키기 위한 반복적인 처리과정을 줄이고, 종래의 양자화 잡음 조절방법에서 발생하는 저비트율 일수록 양자화 잡음을 적절히 분배하지 못하고 양자화 과정을 마침으로써 큰 음질열화를 발생시키는 문제를 해결하기 위한 오디오 인코딩 장치 및 방법을 제공하는데 있다.

【발명의 구성 및 작용】

<31> 상기의 과제를 이루기 위하여 본 발명에 의한 오디오 데이터 인코딩 장치는, 시간 영역의 오디오 신호를 입력받아 주파수 영역의 신호로 변환하는 시간/주파수 매핑부; 상기 변환된 주파수 영역의 오디오 신호를 입력받아 인코딩하고자 하는 오디오 포맷에 대응되는 스펙트럴 처리를 수행하는 스펙트럴 처리부; 상기 변환된 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 마스킹 임계치

계산부; 및 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 양자화 잡음 곡선 조절부를 포함한다.

<32> 상기의 과제를 이루기 위하여 본 발명에 의한 양자화 잡음 분포 조절 장치는, 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 마스킹 임계치 계산부; 및 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 모든 주파수 대역에 대한 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 양자화 잡음 곡선 조절부를 포함한다.

<33> 상기의 과제를 이루기 위하여 본 발명에 의한 오디오 데이터 인코딩 방법은, (a) 시간영역의 오디오 신호를 입력받아 주파수 영역의 신호로 변환하는 단계; (b) 상기 변환된 주파수 영역의 신호에 대해 인코딩하는 오디오 포맷에 맞는 스펙트럴 처리를 수행하는 단계; (c) 상기 변환된 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 단계; 및 (d) 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 단계를 포함한다.

- <34> 상기의 과제를 이루기 위하여 본 발명에 의한 양자화 잡음 분포 조절 방법은, (a) 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 단계; 및 (b) 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 모든 주파수 대역에 대한 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 단계를 포함한다.
- <35> 상기한 과제를 이루기 위하여 본 발명에서는, 상기 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.
- <36> 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 일실시예를 상세히 설명한다.
- <37> 도 3은 본 발명의 오디오 인코딩 장치의 블록도이다.
- <38> 본 발명의 오디오 인코딩 장치는 시간/주파수 매핑부(310), 스펙트럴 처리부(320), 마스킹 임계치 계산부(330), 양자화 잡음 곡선 조절부(340) 및 비트 스트림 생성부(350)를 포함한다.
- <39> 시간/주파수 매핑부(310)는 시간영역 신호를 주파수 영역의 신호로 변환한다. 인코딩하는 포맷에 따라서 시간/주파수 매핑부(310)에서 수행되는 처리가 달라지는데, AAC(Advanced Audio Coding) 포맷 또는 MP3(MPEG-1 layer 3) 포맷으로 인코딩할 때는 MDCT(Modified Discrete Cosine Transform)가 수행된다.

- <40> 스펙트럴 처리부(320)는 주파수 영역의 신호를, 인코딩하는 오디오 포맷에 맞는 스펙트럴 처리를 수행한다. 이러한 스펙트럴 처리의 예로는, TNS(Temporal Noise Shaping), LTP(Long Term Prediction), PNS(Perceptual Noise Substitution), I/C, M/S 등이 있다.
- <41> 마스킹 임계치 계산부(330)는 에너지 분포 곡선 산출부(331), 양자화 잡음 곡선 패턴 추정부(332), 비트조절 초기값 설정부(333)를 포함하며, 입력된 오디오 데이터에 대하여 MDCT를 수행하여 주파수 대역별로 에너지 레벨을 산출하고 심리 음향 모델에 의한 임계 잡음 레벨과 유사한 분포 형태로 근사화시키고, 주파수 밴드별 스케일 팩터 계인을 계산한다.
- <42> 에너지 분포 곡선 산출부(331)는 입력된 오디오 데이터에 대하여 MDCT를 수행하여 주파수 대역별로 에너지 레벨을 산출한다. 양자화 잡음 곡선 패턴 추정부(332)는 산출된 에너지 분포곡선을 기반으로 대역별 이득을 상대적으로 조절하여 양자화 잡음 분포를 설정한다. 비트조절 초기값 설정부(333)는 스케일팩터 밴드 이득만을 결정하는 부분으로, 비트조절 초기값 설정부(333)에서는 글로벌 이득이 초기값을 갖기 때문에 목표 비트율 보다 많은 비트를 사용하는 상태가 된다.
- <43> 도 4a 내지 도 4d는 스케일 팩터 밴드의 에너지를 근사화하는 과정을 설명하기 위한 도면이다.
- <44> 입력된 오디오 데이터에 대하여 MDCT 가 수행되면 도 4a에 도시한 것과 같은 MDCT 라인이 얻어지고, 이를 스케일 팩터 밴드별로 몇개씩 묶어서 나타낸 것이 도 4b이다. 그리고 나서 스케일 팩터 밴드별 에너지를 도 4c의 실선과 같이 조정한다. 양쪽의 스케일 팩터 밴드의 에너지중 어느 한쪽이 자신의 에너지보다 크면 자신의 스케일 팩터 밴드의

에너지를 높이고 그렇지 않으면 그대로 둔다. 이를 수식으로 표현하면 다음 수학적 식 1과 같이 표현된다.

<45> **【수학적 식 1】**
$$M(sfb) = E(sfb) + \alpha |E(sfb-1) - E(sfb)| + \beta |E(sfb+1) - E(sfb)|$$

<46> 여기서, sfb는 스케일팩터 밴드를, M(sfb)은 스케일팩터 밴드별로 근사화된 스케일 팩터 에너지를 의미한다.

<47> 도 4d는 근사화된 스케일 팩터 에너지 곡선이다. 그리고 추정된 M(sfb)을 이용하여 상술한 수학적 식 2에 의해서 스케일팩터 밴드 게인 sfbgain(sfb)을 계산한다.

<48> **【수학적 식 2】**
$$sfbgain(sfb) = \gamma |M(sfb) - E(sfb)|^0$$

<49> 양자화 잡음 곡선 조절부(340)는 이렇게 결정된 주파수 밴드별 스케일 팩터 게인은 고정시킨 채로, 목표 비트율을 만족시키기 위해 모든 주파수 대역에 해당하는 공통이득을 조절하여 양자화 잡음 곡선을 에너지 분포 곡선에 정합시킨다. 정해진 비트율에서 사용될 수 있는 비트수와 비교하여 사용되는 비트수가 정해진 비트율의 비트수보다 적으면 그 비트를 가지고 인코딩을 수행하고, 그렇지 않으면 상술한 양자화 잡음 곡선 조절을 다시 수행한다.

<50> 이렇게, 양자화 잡음의 주파수 대역별로 분포시키는 기준이 되는 임계 잡음 레벨은 심리음향 모델에 의하지 않고, DCT에 의한 주파수 성분만으로 심리음향 모델에 의해 산출된 임계 잡음 레벨과 비슷하고 처리 과정은 간략하게 근사화된 임계 잡음 레벨을 산출한다. 양자화 잡음을 임계 잡음 레벨 이하로 낮추면서 목표 비트율을 만족시키기 위해 글로벌 게인과 스케일 팩터 게인을 반복적으로 많은 횟수의 루프를 수행하지 않고 근사화된 임계 잡음 레벨의 분포와 같은 모양으로 상대적으로 조절한다. 상대적으로 조절된

양자화 잡음의 대역별 비율(스케일팩터 밴드 게인)을 고정시킨채로 목표비트율을 만족시키기 위해 전체 대역에 대한 이득(글로벌 게인)을 조절한다.

<51> 도 5는 본 발명의 오디오 인코딩 방법의 플로우차트이다.

<52> 이제 도 5를 참조하여, 음질 열화를 줄이고 고속으로 오디오 데이터를 인코딩하기 위한 에너지 분포 곡선의 단순 정합 기법에 기반한 MPEG-4 AAC 인코딩 알고리즘을 일실시예로 설명한다.

<53> 시간영역의 오디오 신호를 주파수 영역의 신호로 변환한다(S410). 그리고, 주파수 영역 신호가 가지는 과잉 정보를 줄여주도록 주파수 영역에서의 스펙트럴 처리를 수행한다(S420).

<54> 복잡한 계산을 처리하는 심리음향모델을 통하여 임계잡음 레벨을 구하지 않고 단순히 주파수 영역 신호를 이용하여 주파수 대역별 에너지 레벨을 계산한다(S430). 이때 심리음향모델을 통한 임계 잡음레벨의 형태와 유사하도록 하기 위해 주파수 대역별 에너지 레벨을 근사화한다(S440). 즉, 이웃한 주파수 대역중 어느 하나라도 그 에너지 레벨이 크면 해당 대역의 에너지 레벨을 이웃 대역의 큰 에너지 레벨과의 차이에 대한 일정 비율만큼 에너지 레벨을 증가시킨다. 상세하게는 상술한 수학식1에서 설명한 정도로 증가시킨다.

<55> 다음으로, 조절된 에너지 레벨 분포형태를 통해서 양자화 잡음 분포곡선의 패턴을 추정한다(S450). 입력된 오디오 프레임의 전체 주파수 대역중 가장 큰 에너지 레벨을 갖는 주파수 대역을 찾고, 이를 기준으로 각 주파수 대역별 에너지 레벨과의 차이에 따라 주파수 대역별 이득, 즉, 주파수 대역별 스케일팩터 밴드 이득(scalefactor band gain)

을 결정한다. 이 과정을 통하여 주파수 대역별 양자화 잡음 분포는 에너지 분포를 임계 잡음 형태로 근사화시킨 분포형태를 갖는다.

<56> 목표 비트율에 맞추어 양자화 잡음 분포를 근사화된 에너지 레벨에 정합시키기 위해 비트 조절의 초기치를 결정한다(S460).

<57> (S450) 단계에서 계산된 주파수 대역별 스케일팩터 밴드 이득(scalefactor band gain)를 고정시키고 목표 비트율을 만족시키기 위해 전 대역에 해당하는 공통 이득값을 조절한다(S470). 이렇게 하여 양자화 잡음이 에너지 레벨 분포형태로 근사화된다.

<58> 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

<59> 이제까지 본 발명에 대하여 그 바람직한 실시예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그

와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

【발명의 효과】

- <60> 이상에서 설명한 바와 같이, 본 발명에 따른 오디오 데이터 인코딩 장치 및 방법은 다음과 같은 효과를 제공한다.
- <61> 첫째, 기존의 오디오 인코딩 과정에서 사용되는 심리음향 모델을 직접 사용하지 않고 주파수별 에너지 분포를 통하여 임계잡음 레벨의 상대적인 대역별 분포 형태와 비슷한 분포를 산출함으로써 인코더를 간단하게 구현할 수 있는 효과가 있다.
- <62> 둘째, 기존의 양자화는 제한된 비트수에 대해서 비효율적인 비트 할당을 초래하여 음질 열화에 직접적인 영향을 미치는 것에 비하여, 본 발명은 근사화된 잡음 레벨 분포에 대해 대역별 이득을 비트율 조절보다 먼저 조절함으로써 대역별 양자화 잡음의 상대적 분포를 우선적으로 조절한다. 이렇게 상대적으로 양자화 잡음을 조절한 후 비트율을 조절하는 에너지 분포에 기반한 양자화 잡음 정합 과정을 거치게 되면 기존의 양자화 루프 과정을 통하여 이루어지는 많은 연산량을 획기적으로 감소시킬 수 있으며, 임계 잡음 레벨의 크기 분포와 비슷한 형태로 양자화 잡음의 분포를 얻음으로써 음질 성능을 개선시키는 효과가 있다.
- <63> 세째, 양자화 잡음의 포락선을 DCT를 이용해 근사화된 임계 잡음 레벨의 분포와 절대적으로 만족시키도록 하지 않고 상대적으로 같은 모양을 갖도록 조절하고 나서 비트율을 맞추면, 종래에 주파수 대역에 따라 허용된 임계치를 과도하게 초과하는 현상의 발생을 억제함으로써 오디오 인코딩에서 발생할 수 있는 음질 열화 발생을 현저히 감소시키

는 효과가 있다. 또한 심리음향 모델을 통한 임계잡음 레벨을 산출하는 복잡한 연산과정이 생략되고, 임계 잡음의 절대값에 따라 양자화 잡음을 조절하고 비트율을 맞추는 반복 수행 과정이 생략되어 고속의 오디오 인코딩을 구현할 수 있는 효과가 있다.

【특허청구범위】

【청구항 1】

시간영역의 오디오 신호를 입력받아 주파수 영역의 신호로 변환하는 시간/주파수 매핑부;

상기 변환된 주파수 영역의 오디오 신호를 입력받아 인코딩하고자 하는 오디오 포맷에 대응되는 스펙트럴 처리를 수행하는 스펙트럴 처리부;

상기 변환된 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 마스킹 임계치 계산부; 및

상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 양자화 잡음 곡선 조절부를 포함하는 것을 특징으로 하는 오디오 데이터 인코딩 장치.

【청구항 2】

제1항에 있어서, 상기 시간/주파수 매핑부는

입력된 시간영역의 신호에 대하여 MDCT를 수행하는 것을 특징으로 하는 오디오 데이터 인코딩 장치.

【청구항 3】

제1항에 있어서, 상기 스펙트럴 처리부는

인코딩하고자 하는 오디오 포맷에 따라서 TNS(Temporal Noise Shaping) 또는 LTP(Long Term Prediction) 또는 PNS(Perceptual Noise Substitution)를 수행하는 것을 특징으로 하는 오디오 데이터 인코딩 장치.

【청구항 4】

제1항에 있어서, 상기 마스킹 임계치 계산부는

입력된 오디오 데이터에 대하여 MDCT를 수행하여 주파수 대역별로 에너지 레벨을 산출하는 에너지 분포 곡선 산출부; 및

상기 산출된 에너지 분포곡선을 기반으로 주파수 대역별 이득을 상대적으로 조절하여 양자화 잡음의 분포를 조정하는 양자화 잡음 곡선 패턴 추정부; 및

목표 비트율 보다 많은 비트를 사용할 수 있도록 스케일 팩터 밴드 이득을 결정하는 비트조절 초기값 설정부를 포함하는 것을 특징으로 하는 오디오 데이터 인코딩 장치.

【청구항 5】

제1항에 있어서, 상기 양자화 잡음 곡선 조절부는

정해진 비트율에서 사용될 수 있는 비트수와 비교하여 사용되는 비트수가 정해진 비트율의 비트수보다 적으면 그 비트를 가지고 인코딩을 수행하고, 그렇지 않으면 상기 양자화 잡음 곡선 정합을 다시 수행하는 것을 특징으로 하는 오디오 데이터 인코딩 장치.

【청구항 6】

주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨

곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 마스킹 임계치 계산부; 및

상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 모든 주파수 대역에 대한 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 양자화 잡음 곡선 조절부를 포함하는 것을 특징으로 하는 양자화 잡음 분포 조절 장치.

【청구항 7】

(a) 시간영역의 오디오 신호를 입력받아 주파수 영역의 신호로 변환하는 단계;

(b) 상기 변환된 주파수 영역의 신호에 대해 인코딩하는 오디오 포맷에 맞는 스펙트럴 처리를 수행하는 단계;

(c) 상기 변환된 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 단계; 및

(d) 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 단계를 포함하는 것을 특징으로 하는 오디오 데이터 인코딩 방법.

【청구항 8】

제7항에 있어서, 상기 (c) 단계는

(c1) 상기 변환된 주파수 영역의 신호를 이용하여 주파수 대역별 에너지 레벨을 계산하는 단계;

(c2) 상기 주파수 대역별 에너지 레벨을 근사화하는 단계;

(c3) 상기 근사화된 에너지 레벨 분포형태를 이용하여 양자화 잡음 분포곡선의 패턴을 추정하는 단계;

(c4) 목표 비트율에 맞추어 상기 양자화 잡음 분포곡선을 상기 주파수 대역별 에너지 레벨에 정합시키기 위해 비트 조절의 초기치를 결정하고 주파수 대역별 스케일팩터 밴드 이득을 계산하는 단계; 및

(c5) 상기 주파수 대역별 스케일팩터 밴드 이득을 고정시키고 목표 비트율을 만족시키기 위해 모든 주파수 대역에 대한 공통 이득값을 조절하는 단계를 포함하는 것을 특징으로 하는 오디오 데이터 인코딩 방법.

【청구항 9】

제8항에 있어서, 상기 (c2) 단계는

이웃한 주파수 대역의 신호중 어느 하나라도, 이웃한 주파수 대역의 신호의 에너지 레벨이 크면, 이웃 주파수 대역의 에너지 레벨과 해당 주파수 대역 신호의 에너지 레벨의 차이의 일정 비율만큼, 해당 주파수 대역 신호의 에너지 레벨을 증가시키는 것을 특징으로 하는 오디오 데이터 인코딩 방법.

【청구항 10】

제8항에 있어서, 상기 (c3) 단계는

전체 주파수 대역의 신호중 가장 큰 에너지 레벨을 갖는 주파수 대역의 신호를 찾고, 이를 기준으로 각 주파수 대역별 신호의 에너지 레벨과의 차이에 따라 주파수 대역별 이득을 결정하여 주파수 대역별 양자화 잡음 에너지 분포를 임계 잡음 형태로 근사화시키는 것을 특징으로 하는 오디오 데이터 인코딩 방법.

【청구항 11】

(a) 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 단계; 및

(b) 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 모든 주파수 대역에 대한 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 단계를 포함하는 것을 특징으로 하는 양자화 잡음 분포 조절 방법.

【청구항 12】

(a) 시간영역의 오디오 신호를 입력받아 주파수 영역의 신호로 변환하는 단계;

(b) 상기 변환된 주파수 영역의 신호에 대해 인코딩하는 오디오 포맷에 맞는 스펙트럴 처리를 수행하는 단계;

(c) 상기 변환된 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의

한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 단계; 및

(d) 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 단계를 포함하는 것을 특징으로 하는 오디오 데이터 인코딩 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

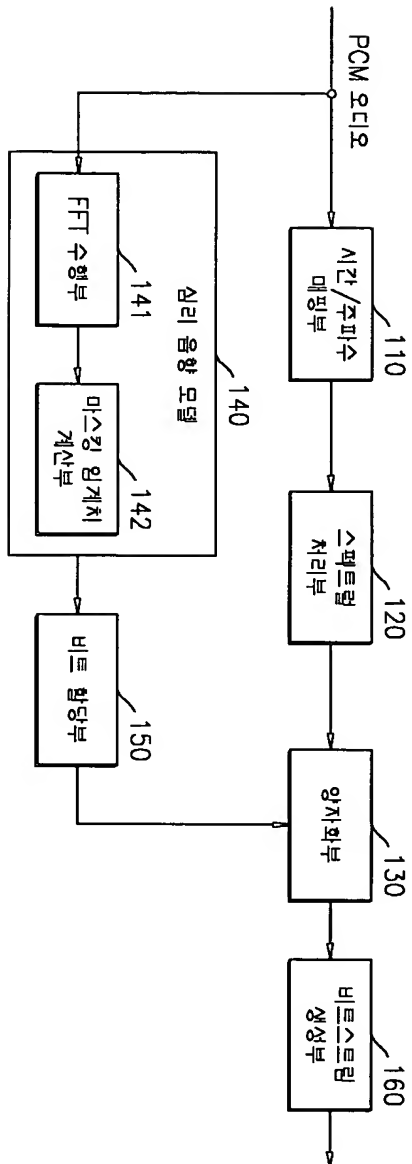
【청구항 13】

(a) 주파수 영역의 오디오 신호를 입력받아 주파수 대역별로 에너지 레벨을 산출하고, 산출한 에너지 레벨의 에너지 분포 곡선이 종래의 심리음향모델에 의한 임계 잡음 레벨 곡선과 유사한 분포 형태가 되도록 근사화시키고 주파수 밴드별 스케일 팩터 밴드 이득을 계산하는 단계; 및

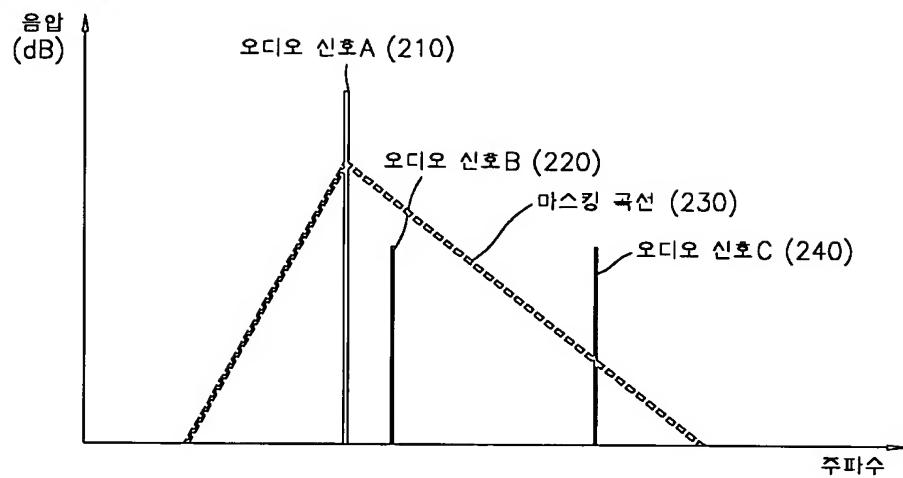
(b) 상기 주파수 밴드별 스케일 팩터 밴드 이득은 고정시킨 채로, 목표 비트율을 만족시키기 위해 모든 주파수 대역에 대한 공통이득을 조절하여 양자화 잡음 곡선을 상기 근사화된 에너지 분포 곡선에 정합시키는 단계를 포함하는 것을 특징으로 하는 양자화 잡음 분포 조절 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

【도면】

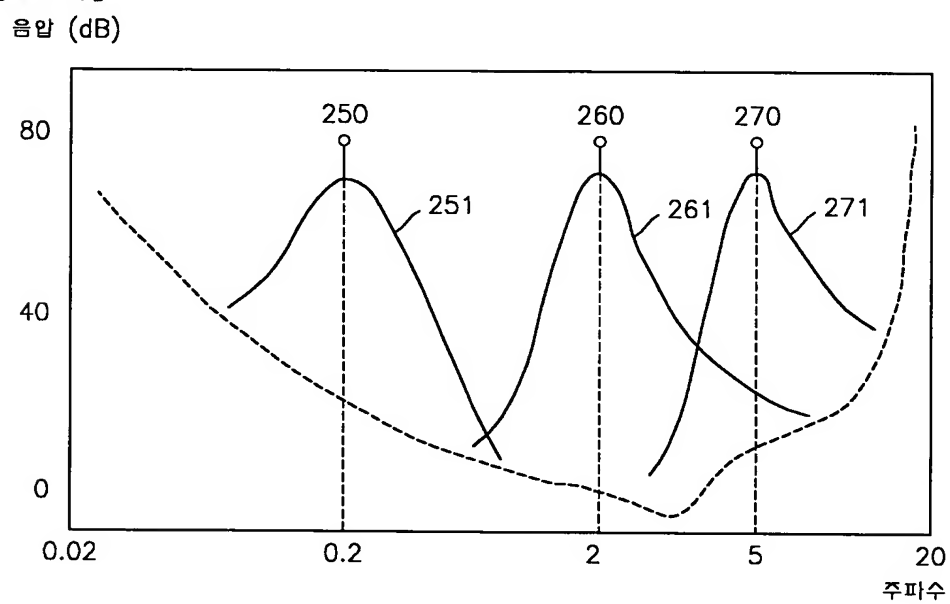
【도 1】



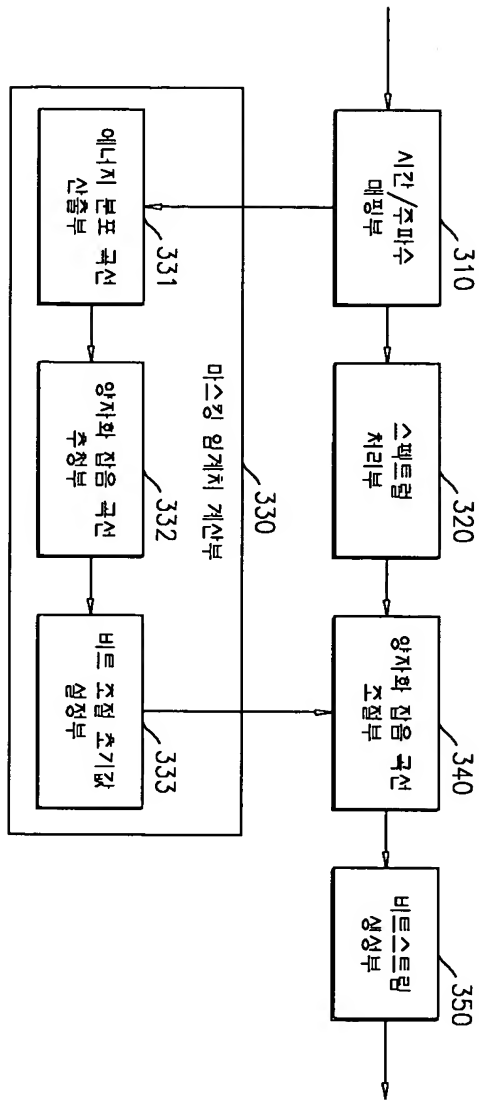
【도 2a】



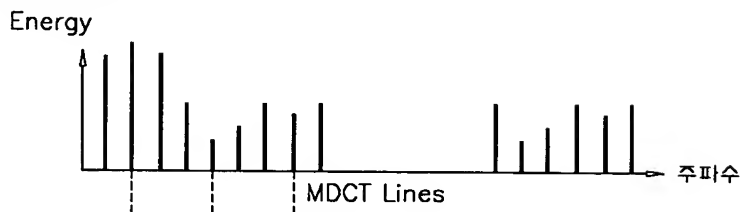
【도 2b】



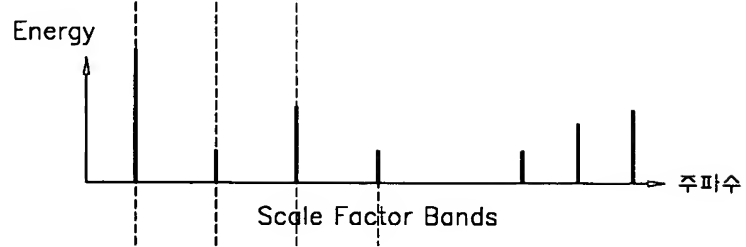
【도 3】



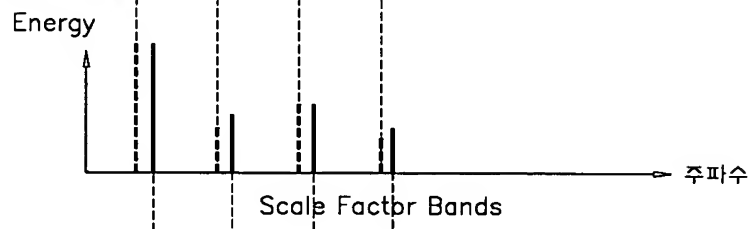
【도 4a】



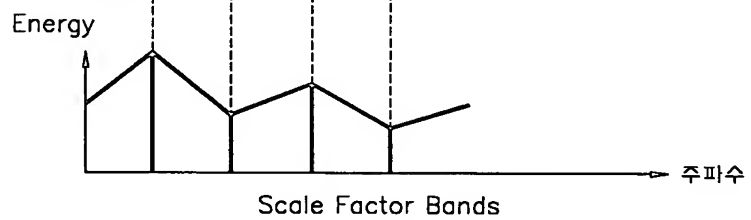
【도 4b】



【도 4c】



【도 4d】



【도 5】

